PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 : C12N 15/00, C12P 21/02 — A1

(11) International Publication Number: WO 89/ 03422

(43) International Publication Date: 20 April 1989 (20.04.89)

(21) International Application Number: PCT/GB88/00831

(22) International Filing Date: 7 October 1988 (07.10.88)

(31) Priority Application Number: 8723661

(32) Priority Date: 8 October 1987 (08.10.87)

(33) Priority Country: GB

(71) Applicant *(for all designated States except US)*: BRITISH BIO-TECHNOLOGY LIMITED [GB/GB]; Brook House, Watlington Road, Cowley, Oxford OX4 5LY (GB).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)* : EDWARDS, Richard, Mark [GB/GB]; 7 Ludlow Drive, Thame, Oxon (GB).

(74) Agents: SHEARD, Andrew, Gregory et al.; Kilburn & Strode, 30 John Street, London WC1N 2DD (GB).

(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), DK, FI, FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), NO, SE (European patent), US.

Published
With international search report.

(54) Title: SYNTHETIC GENE

ATG GAC CCG TCC AAG GAC TCC AAA GCT CAG GTT TCT
GCA GCC GAA GCT GGT ATC ACT GGC ACC TGG TAT AAC
CAA CTG GGG TCG ACT TTC ATT GTG ACC GCT GGT GCG
GAC GGA GCT CTG ACT GGC ACC TAC GAA TCT GCG GTT
GGT AAC GCA GAA TCC CGC TAC GTA CTG ACT GGC CGT
TAT GAC TCT GCA CCT GCC ACC GAT GGC TCT GGT ACC
GCT CTG GGC TGG ACT GTG GCT TGG AAA AAC AAC TAT
CGT AAT GCG CAC AGC GCC ACT ACG TGG TCT GGC CAA (I)
TAC GTT GGC GGT GCT GAG GCT CGT ATC AAC ACT CAG
TGG CTG TTA ACA TCC GGC ACT ACC GAA GCG AAT GCA
TGG AAA TCG ACA CTA GTA GGT CAT GAC ACC TTT ACC
AAA GTT AAG CCT TCT GCT GCT AGC ATT GAT GCT GCC
AAG AAA GCA GGC GTA AAC AAC GGT AAC CCT CTA GAC
GCT GTT CAG CAA TAA

(57) Abstract

Synthetic DNA coding for streptavidin includes sequence (I) and incorporates useful restriction sites at frequent intervals to facilitate the cassette mutagenesis of selected regions. Also included are flanking restriction sites to simplify the incorporation of the gene into any desired expression system.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	ML	Mali
AU	Australia	GA	Gabon	MR	Mauritania
BB	Barbados	GB	United Kingdom	MW	Malawi
BE	Belgium	HU	Hungary	NL	Netherlands
BG	Bulgaria	IT	Italy	NO	Norway
BJ	Benin	JP	Japan	RO	Romania
BR	Brazil	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic	KR	Republic of Korea	SE	Sweden
CG	Congo	LI	Liechtenstein	SN	Senegal
CH	Switzerland	LK	Sri Lanka	SU	Soviet Union
CM	Cameroon	LU	Luxembourg	TD	Chad
DE	Germany, Federal Republic of	MC	Monaco	TG	Togo
DK	Denmark	MG	Madagascar	US	United States of America
FI	Finland				

SYNTHETIC GENE

This invention relates to synthetic genes coding for streptavidin.

Streptavidin is a 60 kD protein isolated from Streptomyces avidinii that binds extremely tightly to the vitamin biotin with a dissociation constant of about 10^{-15} M. Streptavidin is composed of four identical subunits of 15 kD and binds 4 mole biotin per mole of protein. It is structurally related to the protein avidin, a biotin binding protein isolated from chicken egg whites.

Streptavidin has particular utility in that it can be readily conjugated to a range of other proteins such as horseradish peroxidase and alkaline phosphatase. This has allowed the development of sensitive detection systems for biotinylated oligonucleotides, antibodies and other binding proteins with application in the analysis of DNA, RNA, protein and carbohydrates.

The natural gene for streptavidin has been described and was found to encode a protein of 183 amino acids which included a typical leader sequence of 24 amino acids. The N-terminus of the mature protein was found to correspond to residue 25 of the deduced sequence. The amino acid sequence of mature streptavidin deduced from the gene sequence was used as the basis of the design for the synthetic streptavidin genes of the present invention and is shown in Figure 1.

The binding of streptavidin to biotin was described in 1963 by Chaiet et al. (Agents Chemother. 3, 28-32). The cloning of the natural streptavidin gene and its sequence together with the deduced streptavidin amino acid sequence is taught by Argarana et al. (Nucleic Acids Res. 14 1871-1882 (1986)). The production of streptavidin-like polypeptides is disclosed in EP-A-0198015 (Biogen NV).

In order to facilitate the dissection of the structure/function relationships of streptavidin, its incorporation into expression vectors and the production of novel chimeric proteins containing streptavidin functionality an improved novel synthetic gene for the streptavidin produced by Streptomyces avidinii is sought.

It is by no means easy to predict the design of an improved streptavidin gene, since the factors that determine the expressibility of a given DNA sequence are still poorly understood. Furthermore, the utility of the gene in various applications will be influenced by such considerations as codon usage and restriction sites. The present invention relates to a synthetic streptavidin gene which is distinct from the natural streptavidin gene and has advantages in the ease with which it can be modified due to the presence of useful restriction sites.

When synthesising and assembling genes, problems have been encountered when there are inverted or direct repeats greater than eight bases long in the genetic sequence. In addition, areas of unbalanced base

composition such as G/C or A/T rich regions or polypurine/polypyrimidine tracts have been found to lead to inefficient expression. The present invention seeks to overcome or at least alleviate these difficulties.

According to a first aspect of the invention, there is provided DNA coding for streptavidin and having restriction sites for the following enzymes:

EcoRI, NdeI, PstI, SalI, SacI, SnaBI, BspMI, KpnI, FspI, PflMI, DraIII, BalI, BstXI, HpaI, NsiI, SpeI, NheI, BstEII, XbaI, BamHI and HindIII

According to a second aspect of the invention, there is provided DNA including the following sequence:

ATG GAC CCG TCC AAG GAC TCC AAA GCT CAG GTT TCT
GCA GCC GAA GCT GGT ATC ACT GGC ACC TGG TAT AAC
CAA CTG GGG TCG ACT TTC ATT GTG ACC GCT GGT GCG
GAC GGA GCT CTG ACT GGC ACC TAC GAA TCT GCG GTT
GGT AAC GCA GAA TCC CGC TAC GTA CTG ACT GGC CGT
TAT GAC TCT GCA CCT GCC ACC GAT GGC TCT GGT ACC
GCT CTG GGC TGG ACT GTG GCT TGG AAA AAC AAC TAT
CGT AAT GCG CAC AGC GCC ACT ACG TGG TCT GGC CAA
TAC GTT GGC GGT GCT GAG GCT CGT ATC AAC ACT CAG
TGG CTG TTA ACA TCC GGC ACT ACC GAA GCG AAT GCA
TGG AAA TCG ACA CTA GTA GGT CAT GAC ACC TTT ACC
AAA GTT AAG CCT TCT GCT GCT AGC ATT GAT GCT GCC
AAG AAA GCA GGC GTA AAC AAC GGT AAC CCT CTA GAC
GCT GTT CAG CAA TAA

A synthetic streptavidin gene as described above incorporates useful restriction sites at frequent intervals to facilitate the cassette mutagenesis of selected regions. Also included are flanking restriction sites to simplify the incorporation of the gene into any desired expression system.

Codons are those that are favoured by E. coli but it is expected that the DNA would be suitable for expression in other organisms including yeast and mammalian cells.

Certain DNA sequences in accordance with the invention encode the entire mature streptavidin protein together with the required initiator methionine residue but lack the leader sequence present in the natural gene. It is envisaged that the leader sequence appropriate to the expression system of choice would be added to the synthetic gene as required or omitted to allow for intracellular expression of the gene.

According to a third aspect of the invention, there is provided a genetic construct comprising DNA according to the first or second aspect or a fragment thereof. The fragment may comprise at least 10, 20, 30, 40 or 50 nucleotides. A genetic construct in accordance with the third aspect may be a vector, such as a plasmid, cosmid or phage. Alternatively or in addition, a genetic construct in accordance with the third aspect may be a chimeric gene composing all or a fragment of DNA according to the first or second aspect fused to any other sequence of DNA so as to result in a sequence capable of encoding a hybrid protein possessing biotin binding activity.

According to a fourth aspect of the invention, there is provided a process for the preparation of DNA in accordance with the first or second aspect or a genetic construct in accordance with the third aspect, the process comprising coupling successive nucleotides and/or ligating appropriate oligomers.

The invention also relates to other nucleic acid (including RNA) either corresponding to or complementary to DNA in accordance with the first or second aspects.

Preferred embodiments and examples of the invention will now be described. In the following description, reference is made to a number of drawings, in which:

Figure 1 shows the amino acid sequence of streptavidin;

Figure 2 shows a sequence of a streptavidin synthetic gene in accordance with the invention incorporating a summary of restriction sites;

Figure 3 shows a sequence of a synthetic streptavidin gene in accordance with the invention divided into oligonucleotides; and

Figure 4 shows a summary of an assembly procedure used.

EXAMPLE

CONSTRUCTION OF THE GENE

The desired gene sequence was divided into 24

oligodeoxyribonucleotides (oligomers) as depicted in Figure 3. The division was such as to provide 7 base cohesive ends after annealing complementary pairs of oligomers. The end points of the oligomers were chosen to minimise the potential for inappropriate ligation of oligomers at the assembly stage.

The oligomers were synthesised by automated solid phase phosphoramidite chemistry. Following de-blocking and removal from the controlled pore glass support the oligomers were purified on denaturing polyacrylamide gels, further purified by ethanol precipitation and finally dissolved in water prior to estimation of their concentration.

All the oligomers with the exception of the 5′ terminal oligomers BB214 and BB237 were then kinased to provide them with a 5′ phosphate as required for the ligation step. Complementary oligomers were then annealed and the 12 pairs of oligomers ligated together by T4 DNA ligase as depicted in Figure 4. The ligation products were separated on a 2% low gelling temperature (LGT) gel and the band corresponding to the 501/501 bp streptavidin gene duplex was cut out and extracted from the gel. The purified fragment was ligated to EcoRI/HindIII cut DNA of the plasmid vector pUC18. The ligated product was transformed into HW87 and plated on L-agar plates containing 100 mcg ml^{-1} ampicillin. Colonies containing potential clones were then grown up in L-broth containing ampicillin at 100 mcg ml^{-1} and plasmid DNA isolated. Positive clones were identified by direct dideoxy sequence analysis of the plasmid DNA using the 17 base universal primer, a reverse

sequencing primer complementary to the opposite strand on the other side of the polylinker and some of the oligomers employed in the assembly of the gene that served as internal primers. One streptavidin clone was subsequently re-sequenced on both strands to confirm that no mutations were present.

METHODS

All the techniques of genetic manipulation used in the manufacture of this gene are well known to those skilled in the art of genetic engineering. A description of most of the techniques can be found in the laboratory manual entitled Molecular Cloning by T. Maniatis, E.F. Fritsch and J. Sambrook published by Cold Spring Harbor Laboratory, Box 100, New York.

Additional and modified methodologies are detailed below.

1) Oligonucleotide synthesis

The oligonucleotides were synthesised by automated phosphoramidite chemistry using cyanoethyl phosphoramidtes. The methodology is now widely used and has been described (Beaucage, S.L. and Caruthers, M.H. Tetrahedron Letters 24, 245 (1981)).

2) Purification of Oligonucleotides

The oligonucleotides were de-protected and removed from the CPG support by incubation in concentrated NH_3. Typically, 50 mg of CPG carrying 1 micromole of

oligonucleotide was de-protected by incubation for 5 hr at 70° in 600 mcl of concentrated NH_3. The supernatant was transferred to a fresh tube and the oligomer precipitated with 3 volumes of ethanol. Following centrifugation the pellet was dried and resuspended in 1 ml of water. The concentration of crude oligomer was then determined by measuring the absorbance at 260 nm.

For gel purification 10 absorbance units of the crude oligonucleotide were dried down and resuspended in 15 mcl of marker dye (90% de-ionised formamide, 10mM tris, 10 mM borate, 1mM EDTA, 0.1% bromophenol blue). The samples were heated at 90° for 1 minute and then loaded onto a 1.2 mm thick denaturing polyacrylamide gel with 1.6 mm wide slots. The gel was prepared from a stock of 15% acrylamide, 0.6% bisacrylamide and 7M urea in 1 X TBE and was polymerised with 0.1% ammonium persulphate and 0.025% TEMED. The gel was pre-run for 1 hr. The samples were run at 1500 V for 4-5 hr. The bands were visualised by UV shadowing and those corresponding to the full length product cut out and transferred to micro-testubes. The oligomers were eluted from the gel slice by soaking in AGEB (0.5 M ammonium acetate, 0.01 M magnesium acetate and 0.1 % SDS) overnight. The AGEB buffer was then transferred to fresh tubes and the oligomer precipitated with three volumes of ethanol at -70° for 15 min. The precipitate was collected by centrifugation in an Eppendorf microfuge for 10 min, the pellet washed in 80 % ethanol, the purified oligomer dried, redissolved in 1 ml of water and finally filtered through a 0.45 micron micro-filter. The concentration of purified product was measured by determining its absorbance at 260 nm.

3) Kinasing of oligomers

250 pmole of oligomer was dried down and resuspended in 20 mcl kinase buffer (70 mM Tris pH 7.6, 10 mM $MgCl_2$, 1 mM ATP, 0.2 mM spermidine, 0.5 mM dithiothreitol). 10 u of T4 polynucleotide kinase was added and the mixture incubated at 37° for 30 min. The kinase was then inactivated by heating at 85° for 15 min.

4) Annealing

8 mcl of each oligomer was mixed, heated to 90° and then slow cooled to room temperature over a period of an hour.

5) Ligation

5 mcl of each annealed pair of oligomers were mixed and 10 X ligase buffer added to give a final ligase reaction mixture (50 mM Tris pH 7.5, 10 mM $MgCl_2$, 20 mM dithiothreitol, 1 mM ATP. T4 DNA ligase was added at a rate of 100 u per 50 mcl reaction and ligation carried out at 15° for 4 hr.

6) Agarose gel electrophoresis

Ligation products were separated using 2% low gelling temperature agarose gels in 1 X TBE buffer (0.094 M Tris pH 8.3, 0.089 M boric acid, 0.25 mM EDTA) containing 0.5 mcg ml^{-1} ethidium bromide.

7) Isolation of ligation product

The band corresponding to the expected streptavidin gene ligation product was identified by reference to size markers under long wave UV illumination. The band was cut out of the gel and the DNA extracted as follows.

The volume of the gel slice was estimated from its weight and then melted by incubation at 65° for 10 min. The volume of the slice was then made up to 400 mcl with TE (10 mM Tris pH 8.0, 1 mM EDTA) and Na acetate added to a final concentration of 0.3 M. 10 mcg of yeast tRNA was also added as a carrier. The DNA was then subjected to three rounds of extraction with equal volumes of TE equilibrated phenol followed by three extractions with ether that had been saturated with water. The DNA was precipitated with 2 volumes of ethanol, centrifuged for 10 min in a microfuge, the pellet washed in 70 % ethanol and finally dried down. The DNA was taken up in 20 mcl of TE and 2 mcl run on a 2 % agarose gel to estimate the recovery of DNA.

8) Cloning of fragment

0.5 mcg of pUC18 DNA (eg ATCC 37253) was prepared by cleavage with HindIII and BamHI as advised by the suppliers. The digested DNA was run on an 0.8 % LGT gel and the vector band purified as described above.

20 ng of cut vector DNA was then ligated to various quantities of streptavidin DNA ranging from 2 to 20 ng for 4 hr using the ligation buffer described above.

The ligation products were used to transform competent HW87 as has been described. Ampicillin resistant transformants were selected on L-agar plates containing 100 mcg ml^{-1} ampicillin.

9) Isolation of plasmid DNA

Plasmid DNA was prepared from the colonies containing potential streptavidin clones essentially as described (Ish-Horowicz, D., Burke, J.F. Nucleic Acids Research **9** 2989-2998 (1981).

10) Dideoxy sequencing

The protocol used was essentially as has been described (Biggin, M.D., Gibson, T.J., Hong, G.F. P.N.A.S. **80** 3963-3965 (1983)). The method was modified to allow sequencing on plasmid DNA as described (Guo, L-H., Wu, R. Nucleic Acids Research **11** 5521-5540 (1983).

11) Transformation

Transformation was accomplished using standard procedures. The strain used as a recipient in the cloning was HW87 which has the following genotype:

araD139(ara-leu)del7697 (lacIPOZY)del74 galU galK hsdR rpsL srl recA56

Any other standard cloning recipient such as HB101 would be adequate.

CLAIMS

1. DNA coding for streptavidin and having restriction sites for the following enzymes:

EcoRI, NdeI, PstI, SalI, SacI, SnaBI, BspMI, KpnI, FspI, PflMI, DraIII, BalI, BstXI, HpaI, NsiI, SpeI, NheI, BstEII, XbaI, BamHI and HindIII

2. DNA including the following sequence:

ATG GAC CCG TCC AAG GAC TCC AAA GCT CAG GTT TCT
GCA GCC GAA GCT GGT ATC ACT GGC ACC TGG TAT AAC
CAA CTG GGG TCG ACT TTC ATT GTG ACC GCT GGT GCG
GAC GGA GCT CTG ACT GGC ACC TAC GAA TCT GCG GTT
GGT AAC GCA GAA TCC CGC TAC GTA CTG ACT GGC CGT
TAT GAC TCT GCA CCT GCC ACC GAT GGC TCT GGT ACC
GCT CTG GGC TGG ACT GTG GCT TGG AAA AAC AAC TAT
CGT AAT GCG CAC AGC GCC ACT ACG TGG TCT GGC CAA
TAC GTT GGC GGT GCT GAG GCT CGT ATC AAC ACT CAG
TGG CTG TTA ACA TCC GGC ACT ACC GAA GCG AAT GCA
TGG AAA TCG ACA CTA GTA GGT CAT GAC ACC TTT ACC
AAA GTT AAG CCT TCT GCT GCT AGC ATT GAT GCT GCC
AAG AAA GCA GGC GTA AAC AAC GGT AAC CCT CTA GAC
GCT GTT CAG CAA TAA

3. A genetic construct comprising DNA as claimed in claim 1 or 2, or a fragment thereof.

4. A construct as claimed in claim 3, wherein the fragment comprises at least 10 nucleotides.

5. A construct as claimed in claim 3, wherein the fragment comprises at least 20 nucleotides.

6. A construct as claimed in claim 3, wherein the fragment comprises at least 30 nucleotides.

7. A construct as claimed in claim 3, wherein the fragment comprises at least 40 nucleotides.

8. A construct as claimed in claim 3, wherein the fragment comprises at least 50 nucleotides.

9. A construct as claimed in claim 3, which is a vector.

10. A genetic construct as claimed in any one of claims 3 to 9 which is or embodies a chimeric gene composing all or a fragment of DNA according to the first or second aspect fused to any other sequence of DNA so as to result in a sequence capable of encoding a hybrid protein possessing biotin binding activity.

11. A process for the preparation of DNA as claimed in claim 1 or 2 or a genetic construct in accordance with claim 3, the process comprising coupling successive nucleotides and/or ligating appropriate oligomers.

12. DNA substantially as herein described with reference to Figure 2.

FIGURE 1

AMINO ACID SEQUENCE OF STREPTAVIDIN.

```
M R K I V V A A I A V S L T T V S I T A
     -20                 -10

S A S A|D P S K D S K A Q V S A A E A G
    -1|1                 10

I T G T W Y N Q L G S T F I V T A G A D
    20                   30

G A L T G T Y E S A V G N A E S R Y V L
    40                   50

T G R Y D S A P A T D G S G T A L G W T
    60                   70

V A W K N N Y R N A H S A T T W S G Q Y
    80                   90

V G G A E A R I N T Q W L L T S G T T E
   100                  110

A N A W K S T L V G H D T F T K V K P S
   120                  130

A A S I D A A K K A G V N N G N P L D A
   140                  150

V Q Q 159
```

Cleavage site after leader sequence indicated by |.

FIGURE 2

SEQUENCE OF SYNTHETIC STREPTAVIDIN GENE AND LIST OF USEFUL RESTRICTION SITES.

```
          M  D  P  S  K  D  S  K  A  Q  V  S  A  A  E  A  G
GAATTCCATATGGACCCGTCCAAGGACTCCAAAGCTCAGGTTTCTGCAGCCGAAGCTGGT
EcoRI NdeI          StyI                      PstI
CTTAAGGTATACCTGGGCAGGTTCCTGAGGTTTCGAGTCCAAAGACGTCGGCTTCGACCA
        10        20        30        40        50        60

 I  T  G  T  W  Y  N  Q  L  G  S  T  F  I  V  T  A  G  A  D
ATCACTGGCACCTGGTATAACCAACTGGGGTCGACTTTCATTGTGACCGCTGGTGCGGAC
                        SalI
TAGTGACCGTGGACCATATTGGTTGACCCCAGCTGAAAGTAACACTGGCGACCACGCCTG
        70        80        90       100       110       120

 G  A  L  T  G  T  Y  E  S  A  V  G  N  A  E  S  R  Y  V  L
GGAGCTCTGACTGGCACCTACGAATCTGCGGTTGGTAACGCAGAATCCCGCTACGTACTG
 SacI                                                 SnaBI
CCTCGAGACTGACCGTGGATGCTTAGACGCCAACCATTGCGTCTTAGGGCGATGCATGAC
        130       140       150       160       170       180

 T  G  R  Y  D  S  A  P  A  T  D  G  S  G  T  A  L  G  W  T
ACTGGCCGTTATGACTCTGCACCTGCCACCGATGGCTCTGGTACCGCTCTGGGCTGGACT
                    BspMI                KpnI
TGACCGGCAATACTGAGACGTGGACGGTGGCTACCGAGACCATGGCGAGACCCGACCTGA
        190       200       210       220       230       240

 V  A  W  K  N  N  Y  R  N  A  H  S  A  T  T  W  S  G  Q  Y
GTGGCTTGGAAAAACAACTATCGTAATGCGCACAGCGCCACTACGTGGTCTGGCCAATAC
                          FspI                        BstXI
CACCGAACCTTTTTGTTGATAGCATTACGCGTGTCGCGGTGATGCACCAGACCGGTTATG
        250       260       270       280       290       300

 V  G  G  A  E  A  R  I  N  T  Q  W  L  L  T  S  G  T  T  E
GTTGGCGGTGCTGAGGCTCGTATCAACACTCAGTGGCTGTTAACATCCGGCACTACCGAA
                                HpaI
CAACCGCCACGACTCCGAGCATAGTTGTGAGTCACCGACAATTGTAGGCCGTGATGGCTT
        310       320       330       340       350       360

 A  N  A  W  K  S  T  L  V  G  H  D  T  F  T  K  V  K  P  S
GCGAATGCATGGAAATCGACACTAGTAGGTCATGACACCTTTACCAAAGTTAAGCCTTCT
   NsiI           SpeI
CGCTTACGTACCTTTAGCATGTGATCATCCAGTACTGTGGAAATGGTTTCAATTCGGAAGA
        370       380       390       400       410       420

 A  A  S  I  D  A  A  K  K  A  G  V  N  N  G  N  P  L  D  A
GCTGCTAGCATTGATGCTGCCAAGAAAGCAGGCGTAAACAACGGTAACCCTCTAGACGCT
  NheI                                      BstEII  XbaI
CGACGATCGTAACTACGACGGTTCTTTCGTCCGCATTTGTTGCCATTGGGAGATCTGCGA
        430       440       450       460       470       480

 V  Q  Q  *  *
GTTCAGCAATAATAAGGATCCAAGCTT
              BamHI HindIII
CAAGTCGTTATTATTCCTAGGTTCGAA
        490       500
```

FIGURE 2b

SUMMARY OF USEFUL RESTRICTION SITES.

ENZYME	SEQUENCE	POSITION
EcoRI	GAATTC	1
NdeI	CATATG	7
PstI	CTGCAG	44
SalI	GTCGAC	90
SacI	GAGCTC	122
SnaBI	TACGTA	172
BspMI	ACCTGC	201
KpnI	GGTACC	220
FspI	TGCGCA	267
PflMI	CCACTACGTGG	278
DraIII	CACTACGTG	279
BalI	TGGCCA	291
BstXI	CCAATACGTTGG	294
DraIII	CACTCAGTG	327
HpaI	GTTAAC	339
NsiI	ATGCAT	365
SpeI	ACTAGT	381
NheI	GCTAGC	424
BstEII	GGTAACC	463
XbaI	TCTAGA	471
BamHI	GGATCC	496
HinDIII	AAGCTT	502

FIGURE 3

DESIGN OF OLIGOMERS FOR SYNTHETIC STREPTAVIDIN GENE.

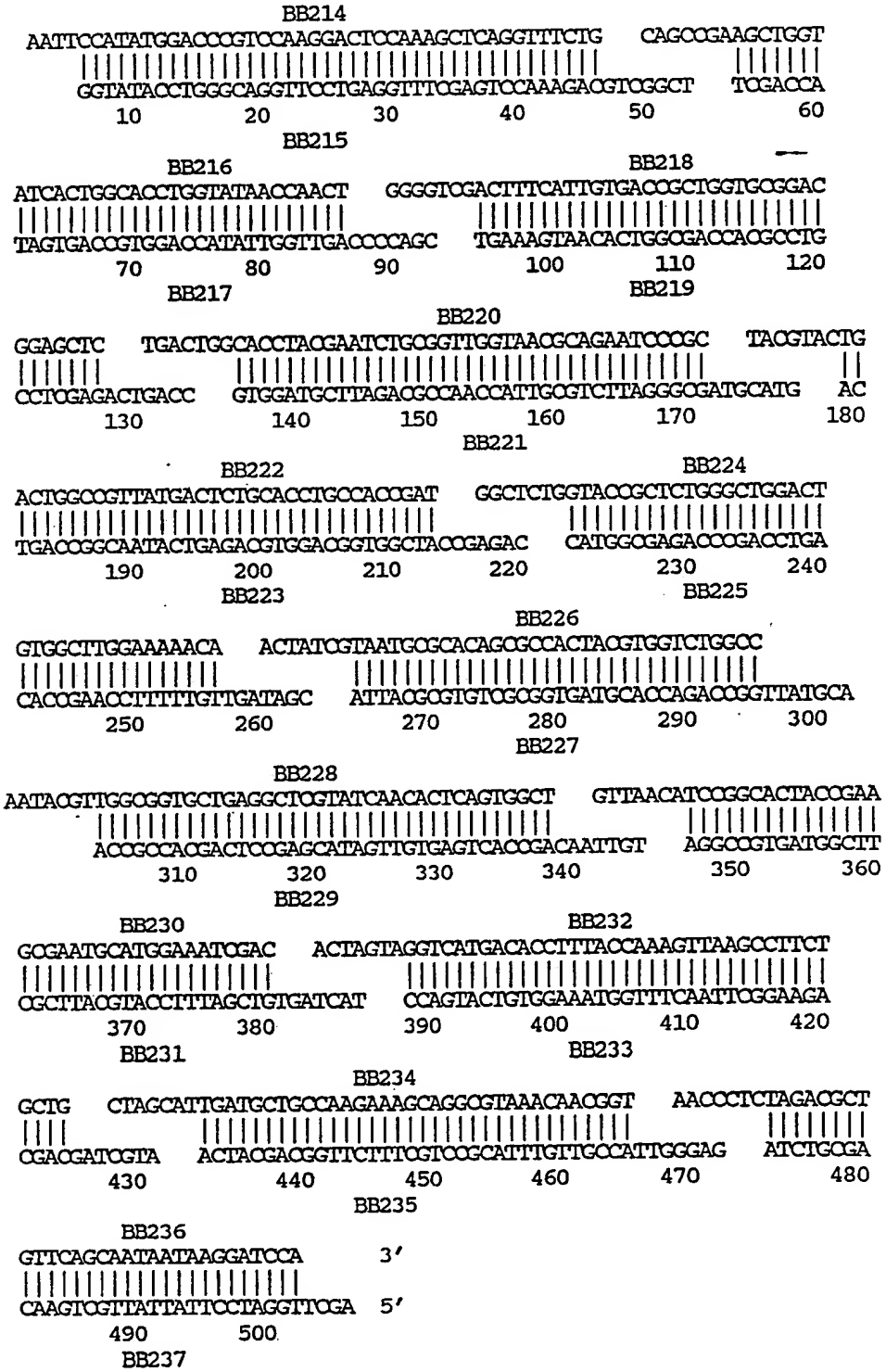
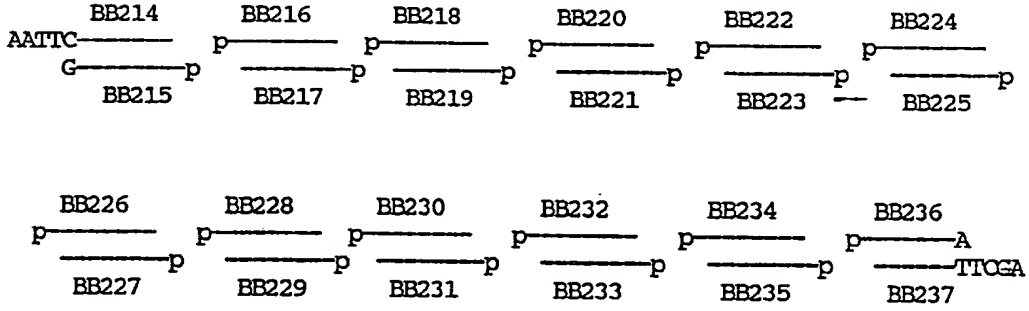


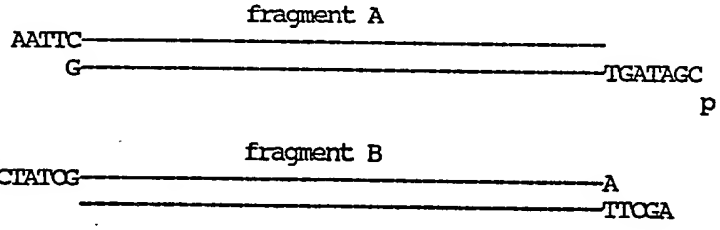
FIGURE 4

SUMMARY OF ASSEMBLY PROCEDURE.

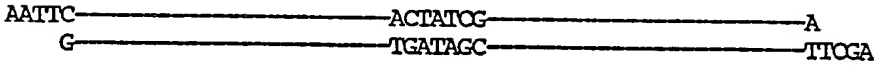
a) kinased oligomers annealed in pairs and mixed in two groups (A & B).



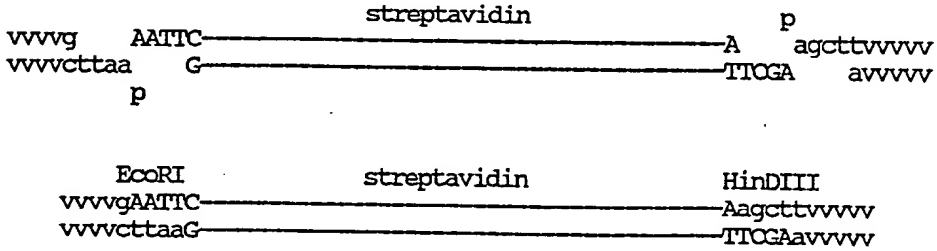
b) oligomers ligated together in two groups. (BB214 and BB237 not kinased to avoid multimerisation.



c) The ligations were checked for the presence of fragment A & B on 2% agarose gels then the ligation reactions were mixed and the reaction allowed to continue to give the final product.



d) Streptavidin gene fragment was isolated on a 2% LGT agarose gel and cloned into EcoRI/HinDIII cut pUC18.



v = vector sequence
p = 5′ phosphates

INTERNATIONAL SEARCH REPORT

International Application No PCT/GB 88/00831

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC[4]: C 12 N 15/00; C 12 P 21/02

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC[4]	C 12 N; C 12 P

Documentation Searched other than Minimum Documentation to the Extent that such Documents are included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category *	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
Y	WO, A, 87/05026 (THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK) 27 August 1987 see the whole document	1-12
Y	WO, A, 86/02077 (BIOGEN N.V.) 10 April 1986 see the whole document cited in the application	1-12
Y	EP, A, 0171024 (HOECHST AG) 12 February 1986 see page 4, lines 5-15; page 5, lines 7-12	1-12
Y	WO, A, 83/04053 (APPLIED MOLECULAR GENETICS) 24 November 1983 see the whole document	1-12

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
16th January 1989	- 2. 02. 89
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	M. VAN MOL

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO. GB 8800831

SA 24658

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on 23/01/89
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A- 8705026	27-08-87	AU-A- 7165287	09-09-87
		EP-A- 0258411	09-03-88
		JP-T- 63502560	29-09-88
WO-A- 8602077	10-04-86	EP-A- 0198015	22-10-86
		JP-T- 62500700	26-03-87
EP-A- 0171024	12-02-86	DE-A- 3429430	20-02-86
		AU-A- 4597785	13-02-86
		JP-A- 61047198	07-03-86
WO-A- 8304053	24-11-83	EP-A- 0108128	16-05-84
		CA-A- 1200515	11-02-86
		US-A- 4695623	22-09-87